DOCUMENT RESUME

ED 063 155 SE 013 769

AUTHOR	Giese, Ronald N.; Redfield, David D.
TITLE	State and National Assessment in Science.
PUB DATE	Apr 72
NOTE	20p.; Paper presented at the National Science Teachers Association Annual Meeting, New York, New York, April 1972
EDRS PRICE	MF-$0.65 HC-$3.29
DESCRIPTORS	Behavioral Objectives; Curriculum Development; *Curriculum Evaluation; *Educational Objectives; Evaluation; *Evaluation Criteria; Science Education; *Secondary School Science; Tests
IDENTIFIERS	*Florida Science Assessment Project

ABSTRACT

The authors describe the development of a curriculum evaluation scheme: the history, literature search, categorization scheme, development of objectives, technical problems, and the future of the model. A sample of the category scheme is included, as well as some performance objectives, definitions of operational verbs used in the objectives, and test items. (CP)

ED 063155

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
OFFICE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY.

STATE AND NATIONAL ASSESSMENT IN SCIENCE

Ronald N. Giese
and
David D. Redfield
Florida Science Assessment Project
Department of Science Education
and
Educational Research Institute
Florida State University
Tallahassee, Florida 32301

A Presentation Given
To
The Association for the Education of Teachers in Science
The National Science Teachers Association
Annual Meeting
New York City

April 7, 1972

SE 013 769

State and National Assessment in Science

The call for assessment and accountability in Education is today's hue and cry at the state, national, and even local levels. National Assessment has attempted to report to the nation about selected knowledge, skills, attitudes, and understandings of young people at several age levels and about the educational progress they are making. It supplies census-like information, not answers. Frequently, citizens - including school administrators, curriculum specialists, science teachers, school board members, and legislators - demand more at state and local levels. Some seek cost justification in streamlining; others, a more meaningful and useful basis for effective planning. All seek improvement in educational practice.

Approaches to the improvement of educational practice are as old as the profession itself, but taxpayer concern may never have been as high. The Florida legislature in 1970 charged the State Board of Education to "adopt comprehensive educational objectives for public education" and appropriated 1.5 million dollars for the creation of the Florida Educational Research and Development Program to do it. The Georgia Educational Accountability and Improvement Act and a companion Teacher Tenure Act will be considered when the Georgia legislature re-convenes. The bill currently before the Georgia House purports "to require the State Board of Education and local school systems to become accountable for the results of public education in relationship to public funds expended, therefore, and to establish objectives and procedures for the systematic improvement of public education." And so it is with state, county, and local schools within many states.

This paper outlines the experience of one group in meeting the challenge of today's call for accountability. This experience is that of the Florida Science Assessment Project - a small staff of four people who with consultant help have acted as a contractor to the Florida State Department of Educational Research and Development Program.

The History

One major objective of the Florida State Department of Education's Educational Research and Development Program has been the development of techniques to assess the results of the various educational programs in Florida. In the academic year 1970-71, the Research and Development Program funded eleven projects each of whose goal was the development of a catalog of objectives and matching criterion-referenced test exercises in a specific subject matter area for a specific grade level. The Florida Science Assessment Project (FSAP), is one of the eleven projects.

In March 1971, a contract was signed with the Florida State University which created FSAP. The contract required FSAP to

1. Plan for and carry out the collection, organization, and revision of a representative set of performance objectives for grades 7, 8, and 9 in the areas of physical science.

2. Develop a categorization scheme for the catalog of objectives of junior high-middle school science.

3. Develop any additional objectives in the physical science area needed in addition to those collected to make a representative set of performance objectives.

4. Write three equivalent criterion-referenced test items for each performance objective collected or developed.

Because of the extensiveness of the tasks, it was decided to augment the permanent staff with several experts who would meet at

a series of short, intensive conferences. These persons with particular expertise were prepared prior to the conferences via correspondence and white papers and brought together for intensive work sessions to obtain the needed interaction among the areas of expertise.

The Literature Search

Pursuant to the task of collecting existing performance objectives, the sources contacted by the Project staff included ERIC, publishers of junior high school science materials, all Florida science supervisors and personnel who had developed objectives, the BSCS, ISCS, IPS, ESCP, and IIS curriculum projects, and National Assessment personnel. In addition, publications of selected Florida and out-of-state school systems were studied.

Upon completion of this search, the staff concluded that no existing set of available objectives met the Research and Development Program's technical requirements. The search did reveal a sizable list of usable proto-objectives from the ISCS Project and also revisions of ISCS objectives made by several Florida counties. However, these proto-objectives did not cover the entire range of needed objectives. The search emphasized the necessity to develop objectives rather than to collect, collate, and revise a set of existing objectives.

While conducting the search for objectives, the staff organized for the second task, the development of the vital categorization scheme. A categorization scheme was needed to define systematically major domains of interest and to provide a perspective in which to develop the catalog of objectives representative of the junior high-middle school science programs and texts being used in Florida.

The Development of a Categorization Scheme

To develop a broad-based scheme, FSAP held a conference at Wakulla Springs, Florida. The participants came to this conference prepared to discuss the question, "What should every student in Florida know and be able to do in science by the end of the ninth grade?" and to develop a categorization scheme based on their answers. The participants were nationally known authors, curriculum developers, and science educators in the area of junior high school science. The concensus of the conference participants was that (1) a content centered classification would be the most useful to the classroom teacher, (2) process skills should be reflected in content sub-categories, and (3) a unified scheme emphasizing pehnomena, natural processes, and interrelationships would be the most desirable. Operating with this concensus, their expertise, and a complete set of the materials of the State adopted junior high school science texts and curricular programs, the participants forged the basic categories and sub-categories of the scheme partially shown in Appendix A.

The Development of Objectives

Having collected a pool of proto-objectives and defined the specific domains to be represented in the catalog, the staff entered into the production stage. A set of proto-objectives were rewritten as objectives using Mager's format. Mager's format, as you may recall, contains statements of (1) the given, (2) question or directive, (3) the respondee, (4) the respondee's behavior, (5) criterion for acceptance or behavior and (6) special limitations. At this point the staff became aware of Gagné's newly developed system for stating an objective. In addition to the six steps in Mager's format, Gagné's system includes

identification of the intellectual skills which the student performs and the content on which he performs these skills in order to exhibit the stated behavior. These additions, it is believed, enhance the communicability of the objectives by (1) stating via one of six explicitly defined verbs the generalizability of the student's knowledge that can be inferred from the correct behavioral response and (2) stating the exact bit of content necessary to the student's correct behavioral response. These verbs as used in the Project are operationally defined in Appendix B. An additional advantage of Gagné's system is that the statement of content ties in well with the content-based classification scheme.

The set of objectives written in Mager's format were rewritten using Gagné's system. This set was then given to the State's Technical Consultant assigned to the Project, who reviewed them to determine if they met the State's criteria. The state's criteria for an objective, as defined in Appendix C, are that it includes a statement of the situation, action, object, limits, and acceptance criterion and that the objective communicates one and only one meaning and requires measurable performance.

Assured that the objectives submitted met the State's requirements, the staff developed the remainder of the proto-objectives into objectives. Also the state-adopted junior high school science texts were again reviewed as sources of proto-objectives to enhance the representativeness of the catalog.

Concurrently, the staff wrote three equivalent criterion-referenced test exercises for each of a pilot set of objectives.

These test exercises were also reviewed by the State's Technical Consultant. They were examined for (1) congruence, (2) integrity, (3) comprehensibility, (4) objectivity, and (5) equivalence, as defined in Appendix C.

At a writing conference, junior high school science teachers and the staff wrote the three equivalent test exercises for the remainder of the objectives. Classroom teachers were employed as item writers because of their expertise in writing test questions using a vocabulary and style intelligible to students. Overall, the experience of the staff was that good criterion-referenced test exercises and performance objectives were best developed together.

After many in-house revisions, the catalog of objectives and exercises was submitted to a panel of subject matter specialists, science educators, county supervisors, and classroom teachers. Each was asked for a critical review of the product from the viewpoint of his area of expertise. The resulting suggestions for revisions were incorporated into the final product. A sample of the FSAP objectives, which reflects the various intellectual skills, are included in Appendix D.

The Technical Problems

Three technical areas of concern to the Project staff are the communicability of the objectives, the congruency of each objective and its associated criterion-referenced test exercises, and the comprehensibility of the test exercises by students.

Communicability involves the question, "Is an objective consistently interpreted in defining a unique task?" It is important that this be

the case if there is to be communication between individual teachers and between the individual teacher and the school system or State about the expected or achieved outcomes of instruction. The two primary causes for lack of communicability found by the project are inadequately defined technical terms which do not communicate their meaning and the style resulting from the State's criteria and objective format.

The lack of congruency between an objective and its test exercises usually takes one of three forms. In its most obvious form, the tasks as stated in the objective and the exercises both relate to the same content but require different behaviors. For example, the objective may state, "When asked to state a definition for the term compound . . ." but the exercises direct the student to select the definition from a set of stated options. A less obvious form involves the case in which the objective and the exercise require the same or similar processes but to different degrees resulting in different behaviors. An example would be the objective which states, "Given an orchid and asked to identify its common group name . . ." but an exercise in which the student is given an orchid and asked to identify which of the following is the common group name for the plant:

1. pine
2. camellia
3. maple
4. rose
5. none of these

If the student answers "none of these," he gets the exercise correct, not because he knows what the plant is, but because he knows what it is not.

In the most subtle form the objective and the exercise seem congruent when analyzed in vacuo. The problem is recognized only when one talks to students and finds that most of the students exhibit the specified behavior not by using the intellectual skill and content stated, but rather by short-circuiting the objective process via another path. This would occur, for example, in the case in which the objective and exercise asked the student to generate an explanation for a phenomenon which is taught as part of the course and therefore his explanation is actually simple recall.

The unambiguous written communication of a task via test exercises is often difficult even when teachers and student are in daily contact. Thus, notwithstanding the expertise of excellent teachers who aided in writing the exercises, one wonders how many of the exercises have shortcomings such as the following: (1) words which can be misunderstood because they have both technical and non-technical meanings, such as *boiling, salt,* and *concentration,* (2) the use of vocabulary or sentence structure which is too sophisticated for most students, and (3) the omission of simple but key directions which are unnecessary or redundant for adults, but not for students.

The Future

In the immediate future, FSAP is under contract to develop catalogs for the earth, space, and biological sciences similar to the catalog developed for physical science. Again, each of the objectives developed will be accompanied by three equivalent criterion-referenced test items and will meet the same criteria. In addition, each objective will be accompanied by a pre-objective. This pre-objective

will state briefly the general situation and action required by the student. Pre-objectives should serve the teacher as an annotated index in selecting objectives and will be a helpful intermediate step in the objective writing process. The development of pre-objectives in elementary school science will be carried out simultaneously by a group in Escambia County.

State plans call for the catalog of objectives and exercises to be submitted to field testing. This testing will involve a general field test followed by an item analysis and a series of small group interviews with some of the students and teachers involved. These data should provide a basis for determining if the current catalog materials have technical utility; that is, that the objectives are communicable and the exercises congruent and comprehensible.

In addition, it could be profitable to expand and organize the objectives into hierarchies so that a student's progress or problems could be assessed accurately, and the teacher could truly "begin or continue from where the students are."

Related and also very profitable to the classroom teacher would be the development of possible sequences of objectives for students with different learning styles and paces. The sequences could permit such differences in learning styles and paces as concrete or formal, fast or slow, verbal or manipulative.

The Uses of Objectives

Awareness of the State's future development of catalogs of objectives and their accompanying test exercises often prompts the following question: "By whom are these catalogs to be used and for what purposes?"

The major projected users and uses include the following three groups.

1. User: Teacher

   Use: In class assessment

For this purpose, the individual teacher would select objectives to determine the following. What skills and understanding does Johnny have upon which to build? What can Johnny do as a result of the instruction? Which instructional techniques are successful? Which need changing?

2. User: Individual Schools and County School Systems

   Use: Program Assessment

For this purpose, the objectives would be selected at school or county level to determine the following. Does program X produce the educational outcomes it claims? Does program X or program Y produce the desired educational outcomes more economically? What students profit most by a given approach? (Note: Each program must be evaluated in terms of the catalog objectives which are appropriate to it, not on the total catalog.)

3. User: State

   Use: State Assessment

For this purpose, the objective selection would be done at the State level by the State subject matter supervisor in conjunction with county supervisors and teachers. The number of objectives would be but a small percentage of the categories and would represent minimum levels of achievement.

Appendix A

Sample of the Category Scheme Used by FSAP

1.0 The student will be able to identify and describe the characteristic processes and assumptions that permeate and define the domain of science.

2.0 The student will be able to demonstrate that substances have different physical and chemical characteristics and be able to identify some substances by determining their characteristics.

3.0 The student will be able to apply the rules of energy, interacting with matter, in accounting for all forms of motion, work, heat, light, sound, electricity, and magnetism.

4.0 The student will be able to identify and apply four basic criteria for recognizing chemical changes in matter interactions. They include that mass is conserved, that there is a definite relation between interacting substances and the products formed, that there is an exchange of energy, and that these interactions can be exemplified under general types.

5.0 The student will be able to describe and interpret characteristics of various forms of energy using plausible models when necessary.

5.1 Heat, its nature and explanation
- 5.11 Kinetic theory -- a model for heat
- 5.12 Effects of heat -- energy states
  solids, liquids, gases, expansion
- 5.13 Pressure -- temperature relationships (qualitative)
- 5.14 Transfer of heat energy -- conduction, convection, radiation
- 5.15 Measurement

5.2 Sound, its nature and explanation
- 5.21 Sound as motion
- 5.22 Characteristics -- speed in media, amplitude, frequency
- 5.23 Application to music

5.3 Light, its nature and explanation
- 5.31 Radiant energy that can do work
- 5.32 Simple optical characteristics -- reflection, absorption, refraction, color

5.4 Electricity as energy that can do work
- 5.41 Electrostatics
  - Concept of charge and its storage
  - Interaction between charges
- 5.42 Current electricity
  - Circuitry, series, parallel, resistance, fuses
  - Source in chemical potential energy or magnetic induction
  - Measurement: volts, amperes, resistance relationships
  - A.C. and D.C. current -- measurement and generation

6.0 The student will be able to describe and apply the general scheme of organizing matter into basic units and aggregates, which have a mechanism for holding units and aggregates in various combinations that are electrical in nature, and which have the physical states that are the result of energy levels inherent in the matter.

7.0 The student will be able to describe or identify in terms of their properties major materials, groupings, or systems of the lithosphere, hydrosphere, atmosphere, exosphere, and biosphere systems.

8.0 The student will be able to detect and classify major processes within and among the five systems.

9.0 The student will be able to make interpretations and predictions from observations in terms of processes within and among the five systems.

10.0 The student will be able to identify the ways man has interacted with the five systems and the ways these interactions have affected both the systems and man.

Appendix B

Operational Definitions of Gagné's Verbs

Characteristic to each performance objective in the FSAP catalog is a verb which denotes one of six intellectual skills, a cognitive strategy or the possession of an attitude. Many of these verbs were formally defined by Robert Gagné.[8] However, their particular application to junior high-middle school science required some modification. The verbs defined for the *Catalog of Performance Objectives* in terms of the way that the student's response to the objectives should be evaluated.

Recalls: The student can recite or write a statement by rote. Successful accomplishment of the task involved in objectives using this verb shows only that the student knows the proper sequence of words to answer a question. It implies nothing about the ability of the student to apply the information stated in the answer.

Manipulates: The student can do a physical task as a rote procedure. Successful accomplishment of the task involved in objectives using this verb shows that the student can maneuver equipment and materials according to a given procedure. It does not mean that the student would or could apply the procedure to a problem.

Discriminates: The student can state whether two objects are the same or different. Successful completion of a task involving this verb means that the student has the ability to separate identical from non-identical objects. It does not imply that the student can cite aspects of the two objects that are the same or different.

---

[8] Robert Gagné, *Defining Objectives for Six Varieties of Learning* (Washington, D.C.: American Educational Research Association, 1971).

Identifies: The student can select an object by its characteristics when given its name. Successful completion of the task of an objective involving this verb means that the student has learned enough of the properties of an object (including its name) to select it from a group of objects. It does not mean that the student understands how or why the object functions or that he would use it properly in a situation in which it is needed.

Classifies: The student can use a defined concept to put an entity into a class. Successful completion of a task involving this verb means that the student has mastered an abstraction or defined concept to a point where he can use it to place abstract or concrete entities into designated categories. Again it does not mean that the student understands how or why the entity functions or that he could make appropriate application of it.

Demonstrates: The student can apply a specific concept, rule, or procedure to a specific task. Successful completion of the task involved in an objective using this verb shows that the student has learned enough of the concept to be able to use it in a specific type of task. It does not imply that he has learned all aspects of the concept so that he could use it in any type of task to which it could apply. Therefore, in many cases there is more than one *demonstrates* objective for a concept to help in assessing the breadth of the student's ability to apply the concept.

Generates: The student can synthesize two or more rules, definitions, or concepts to solve a specific problem. Successful completion of the task involved in an objective using this verb shows that the student has

learned a set of concepts well enough to put them in combinations to do a task. The concepts involved in these objectives should be assessed separately using objectives involving subordinate intellectual skills.

Originates: The student proposes and solves original problems. Successful completion of the task involved in an objective using this verb implies that the student knows the material so well he can identify problems and bring the proper concept and procedures to it in order to solve it.

Choose: The student decides to behave in a certain way in a given situation. Successful completion of the task involved in one situation in an objective using this verb does not imply that the student would show the same attitude in another situation. Since class materials and situations change from day to day, it is suggested that the objective be assessed during a week or two by means of spot checks to establish if the student shows a given attitude in a range of situations.

Appendix C

PRE-OBJECTIVES

A pre-objective is a statement in measurable terms of a behavior to be exhibited by the learner. It possesses each of the following characteristics:

1. Situation - The situation confronting the learner is clearly stated.

2. Action - The action required of the learner is identified.

3. Object - The object on which the learner is to operate (i.e., the object of the action) is clearly stated.

4. Limits - The particular limits of the activity expected of the learner are stated.

5. Measurability - An observable performance is generated by the stated action.

PERFORMANCE OBJECTIVES

Definition: A performance objective is a statement in precise, measurable terms of a particular behavior to be exhibited by the learner under specified conditions. It possesses each of the elements or characteristics specified below:

1. Situation - The situation confronting the learner is clearly specified, including the mode in which stimuli are to be presented.

2. Action - The action required of the learner is unambiguously defined, including the mode in which responses are to be made.

3. Object - The object on which the learner is to operate (i.e., the object of the action) is clearly stated.

4. Limits - The particular limits of the activity expected of the learner are specified.

5. Measurability: The specified action involves an observable rather than an inferred response.

6. Communicability - The objective is so stated that one, and only one, interpretation of the objective is reasonably possible.

7. Criterion - The degree of proficiency required as evidence of accomplishment by a student of the objective is indicated. (The

criterion may be indicated implicitly or explicitly. If implicit, 100% accuracy is effectively designated. If explicit, may be appended parenthetically to the statement of the objective.)

CRITERION-REFERENCED TEST ITEMS (EXERCISES)

Definition: A criterion-referenced test item is an item based upon a performance objective and is designed to allow the determination of whether or not the learner has accomplished the objective. It possesses each of the characteristics specified below:

1. Congruence - The task specified in the item corresponds directly to the performance specified in the objective, including the situation, action, object and limits.

2. Comprehensibility - The item-specified task is so stated or portrayed that the learner clearly understands what is expected of him.

3. Objectivity - The item is stated in such a way that all competent observers (evaluators) can make a clear and unequivocal decision as to whether or not the learner has demonstrated an acceptable performance.

4. Integrity - The item should be structured in such a way that an acceptable response to the item constitutes sufficient evidence, in and of itself, that the learner has accomplished the corresponding objective. (Integrity must pertain to an item whenever feasible. Responsibility for demonstrating infeasibility rests with the producer of the item.)

5. Equivalence - If two or more items corresponds to a single objective, each item in the set should be a true alternate, in that a student who passes (or fails) one item on a given occasion would be expected to pass (or fail) any other item in the set.

Appendix D

Examples of Objectives and Test Exercises

5.15 1 (19)
Manipulates a thermometer to measure temperature.
Given a thermometer in Celsius or Fahrenheit degrees and two liquids at different temperatures and asked to measure and report their temperatures, the student manipulates the thermometer to measure the temperatures of the given substances by reporting the temperatures within the range of + 1°C or + 1°F of the teacher obtained value.
Item A: Get two containers of liquid and a thermometer from your teacher. Measure and report the temperature of the two liquids.

5.31 1 (8)
Recalls the velocity (speed) of light in air.
When asked to state the velocity (speed) of light in air using the proper units for the number given, the student recalls the velocity of light in air as either 186,000 miles per second or 3×10^8 meters/seconds or an equivalent expression by so stating.
Item A: What is the velocity (speed) of light ~~in air?~~ Be sure to state the units for your answer.

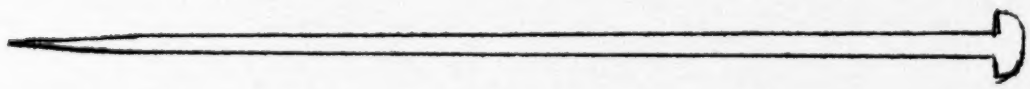
5.21 3 (2)
Identifies the vocal chords as the vibrating objects which produce sound when man talks.
Given a list of human organs of the head-thoracic region and asked to identify the vibrating objects which produce sound when man talks, the student identifies the vocal chords as the vibrating objects which produce sound when man talks by selecting "vocal chords."
Item A: Check the correct term. ~~What vibrates when~~ a person talks?
_____ trachea
_____ mouthpiece
_____ lungs
_(✓)_ vocal chords

5.22 4 (7)
Classifies the comparable characteristics of color in light to pitch in sound.
Given four of the terms from the list -- *pitch, volume, velocity, timbre, amplitude,* and *color* -- and asked to select the terms comparable in sound with the term *color* in light, the student classifies pitch to be the comparable characteristic to sound that color is to light by selecting "pitch."
Item A: ~~Place a check~~ in front of th item which correctly completes the following statement. Color is to light as _____ is to sound.
_(✓)_ pitch _____ volume _____ velocity _____ color

5.32 5 (21)
Demonstrates an application of the rule for computing the length of an object from its picture and its stated magnification.

Given a diagram and the amount of magnification of an object, a device for linear measurement and asked to compute the size of the actual object, the student demonstrates an application of the rule for computing the actual size by multiplying the reciprocal of the magnification by the measured size of the diagram of the object and reporting his answer in the proper units to within + 5%.

Item A: This is a picture of a pin magnified four times. What is the length of the actual pin? Express your answer in the proper units.



(Answers: 1 1/8", 2.9 cm.)

5.42 6 (3)
Generates a procedure to determine which of two batteries can exert more electrical energy.

Given a list of equipment including two "D" size batteries, three test leads, a #222 bulb and socket, string, a timer, a small motor, and eight sinkers and asked to diagram the experimental system he would use to determine which battery has more stored energy and to explain the criterion for his determination, the student generates a procedure by diagraming a completed circuit for any one of the following systems and stating how he would use his system to determine which battery has more influence.

(a) Find out which battery can keep a bulb burning longer.
(b) Find out which battery can cause a motor to move some object farther.
(c) Find out which battery will keep a motor running longer.
(d) Find out which battery can cause a motor to move some object faster.
(e) Find out which battery can cause a motor to move the heavier load.

Teacher's Note: List the following equipment at the beginning of all three items.

2 "D" size batteries
3 test leads
1 #222 bulb and socket
1 timer
1 small motor
8 sinkers
string

Item A: Using any of this equipment, draw a diagram for an experiment to determine which battery exerts more electrical energy. Describe the experiment. State how the results would indicate which battery has more electrical energy.